

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

LẠI HỮU DƯƠNG

**DỰ BÁO CHUỖI THỜI GIAN MỜ DỰA TRÊN
ĐẠI SỐ GIA TỬ VỚI MÔ HÌNH NGỮ NGHĨA
ĐỊNH LƯỢNG TỐI ƯU VÀ ỨNG DỤNG**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Thái Nguyên – 2017

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

LẠI HỮU DƯƠNG

**DỰ BÁO CHUỖI THỜI GIAN MỜ DỰA TRÊN
ĐẠI SỐ GIA TỬ VỚI MÔ HÌNH NGỮ NGHĨA
ĐỊNH LƯỢNG TỐI ƯU VÀ ỨNG DỤNG**

Chuyên ngành: Khoa học máy tính

Mã số: 60 48 01 01

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Người hướng dẫn khoa học: TS. NGUYỄN DUY MINH

(Luận văn đã được sửa theo góp ý của hội đồng bảo vệ thử)

Thái Nguyên - 2017

LỜI CẢM ƠN

Em xin gửi lời cảm ơn chân thành đến người hướng dẫn khoa học - TS. Nguyễn Duy Minh, người đã định hướng và nhiệt tình hướng dẫn, giúp đỡ em trong quá trình làm luận văn.

Em xin gửi lời biết ơn sâu sắc đến quý thầy cô giáo trường Đại học Công nghệ thông tin và Truyền thông Thái Nguyên; Viện công nghệ thông tin thuộc Viện hàn lâm Khoa học và Công nghệ Việt Nam đã truyền đạt những kiến thức và kinh nghiệm quý báu cho chúng em trong thời gian học tập.

Xin chân thành cảm ơn các bạn bè, đồng nghiệp, ban cán sự và các học viên lớp cao học CK14, những người thân trong gia đình đã động viên, chia sẻ, tạo điều kiện giúp đỡ trong suốt quá trình học tập và làm luận văn.

Mặc dù đã rất nỗ lực, cố gắng nhưng chắc chắn luận văn của em vẫn còn nhiều thiếu sót. Rất mong nhận được những ý kiến đóng góp, chia sẻ của quý thầy cô và các bạn.

Xin chân thành cảm ơn!

Thái Nguyên, tháng 6 năm 2017

Tác giả

Lại Hữu Dương

LỜI CAM ĐOAN

Tôi xin cam đoan luận văn “Dự báo chuỗi thời gian mờ dựa trên đại số gia tử với mô hình ngữ nghĩa định lượng tối ưu và ứng dụng” của tôi được thực hiện dưới sự hướng dẫn khoa học của TS. Nguyễn Duy Minh, số liệu và kết quả nghiên cứu trong luận văn này hoàn toàn trung thực và chưa sử dụng để bảo vệ một công trình khoa học nào, các thông tin, tài liệu trích dẫn trong luận văn đã được chỉ rõ nguồn gốc ở phía cuối luận văn.

Mọi sự giúp đỡ cho việc hoàn thành luận văn đều đã được cảm ơn. Nếu có phát hiện nào về sự gian lận trong sao chép tài liệu, công trình nghiên cứu của tác giả khác mà không được ghi rõ trong tài liệu tham khảo, tôi hoàn toàn chịu trách nhiệm.

Thái Nguyên, tháng 6 năm 2017

Tác giả

Lại Hữu Dương

MỤC LỤC

LỜI CẢM ƠN	i
LỜI CAM ĐOAN	iv
MỤC LỤC.....	v
MỤC LỤC HÌNH ẢNH.....	vii
MỤC LỤC BẢNG	viii
MỞ ĐẦU.....	1
CHƯƠNG 1: CÁC KIẾN THỨC CƠ BẢN	5
1.1. Những vấn đề cơ sở của lý thuyết tập mờ và logic mờ	5
1.1.1. Lý thuyết tập mờ	5
1.1.2. Định nghĩa logic mờ.....	6
1.1.3. Các phép toán trên tập mờ.....	7
1.2. Chuỗi thời gian mờ	11
1.3. Đại số gia tử và một số tính chất	14
1.3.1. Đại số gia tử của biến ngôn ngữ.....	14
1.3.2. Độ đo tính mờ và ánh xạ định lượng ngữ nghĩa.....	17
1.4. Bài toán tối ưu và giải thuật di truyền	23
1.4.1. Bài toán tối ưu	23
1.4.2. Giải thuật di truyền.....	24
1.5. Kết luận chương 1.....	28
CHƯƠNG 2: MÔ HÌNH DỰ BÁO CHUỖI THỜI GIAN MỜ DỰA TRÊN ĐẠI SỐ GIA TỬ (ĐSGT).....	29
2.1. Một số mô hình chuỗi thời gian mờ.....	29
2.1.1. Thuật toán của Song và Chissom	29
2.1.2. Thuật toán của Chen.....	30
2.2. Mô hình tính toán và thuật toán dự báo mờ dựa trên đại số gia tử với mô hình ngữ nghĩa định lượng tối ưu.	32
2.2.1. Mô hình dự báo mờ sử dụng đại số gia tử.....	32
2.2.2. Thuật toán dự báo mờ dựa trên đại số gia tử với mô hình ngữ nghĩa định lượng tối ưu	34

2.3. Kết luận chương 2.....	40
CHƯƠNG 3: ỨNG DỤNG MÔ HÌNH DỰ BÁO DỰA TRÊN ĐẠI SỐ GIA TỬ VỚI THAM SỐ NGỮ NGHĨA ĐỊNH LƯỢNG TỐI ƯU	41
3.1. Xây dựng mô hình dự báo chuỗi thời gian mờ.....	41
3.1.1. Mô hình dự báo sinh viên nhập học của trường đại học Alabama của Song và Chissom	41
3.1.2. Mô hình dự báo chuỗi thời gian mờ sinh viên nhập học của trường đại học Alabama của Chen.....	47
3.2. Ứng dụng mô hình dự báo dựa trên đại số gia tử với tham số ngữ nghĩa định lượng tối ưu.....	55
3.2.1. Mô hình dự báo mờ dựa trên đại số gia tử	55
3.2.2. Mô hình dự báo mờ dựa trên Đại số gia tử với ngữ nghĩa định lượng tối ưu	63
3.3. Kết luận chương 3.....	70
PHẦN 3: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN	71
TÀI LIỆU THAM KHẢO.....	72

MỤC LỤC HÌNH ẢNH

Hình 1. 1: Giao của hai tập mờ	8
Hình 1.2: Phép hợp của hai tập mờ	9
Hình 1.3. Minh họa lai ghép.....	26
Hình 3.1: Số sinh viên nhập học thực tế và số sinh viên nhập học dự báo....	47
Hình 3.2: Dữ liệu tuyển sinh thực tế và dữ liệu tuyển sinh dự báo	55

MỤC LỤC BẢNG

Bảng 1.1: Các cặp T - chuẩn và T - đối chuẩn.....	9
Bảng 1.2: Một số phép kéo theo mờ thông dụng	10
Bảng 1.3: Ví dụ về tính âm dương giữa các gia tử	15
Bảng 3.1: Chuyển đổi các giá trị lịch sử thành giá trị ngôn ngữ	43
Bảng 3.2: Xác định các quan hệ thành viên.....	45
Bảng 3.3: Mờ hóa chuỗi dữ liệu	49
Bảng 3.4: Quan hệ logic mờ của dữ liệu tuyển sinh	49
Bảng 3.5: Các nhóm quan hệ logic mờ	50
Bảng 3.6: Bảng so sánh các phương án dự báo.....	54
Bảng 3.7: Số sinh viên nhập học tại trường đại học Alabama từ 1971 đến 1992.....	56
Bảng 3.8: Giá trị đầu và giá trị cuối của các khoảng giải nghĩa được chọn	61
Bảng 3.9: Kết quả tính toán dự báo tối ưu số sinh viên nhập học tại trường đại học Alabama từ 1971 đến 1992 theo tiếp cận ĐSGT	63
Bảng 3.10: So sánh các phương pháp dự báo với 7 khoảng chia	67
Bảng 3.11: So sánh các kết quả mô hình dự báo tối ưu theo tiếp cận ĐSGT và các kết quả mô hình dự báo cải tiến khác.....	69

MỞ ĐẦU

Trong những năm gần đây, có rất nhiều tác giả trên thế giới quan tâm nghiên cứu mô hình dự báo chuỗi thời gian mờ. Nhiều nghiên cứu ứng dụng dự báo có giá trị thực tế đã được thực hiện trên cơ sở phương pháp luận dự báo theo mô hình chuỗi thời gian mờ nêu trên. Vì vậy cho đến nay, mô hình dự báo chuỗi thời gian mờ luôn được nhiều chuyên gia trên thế giới và Việt Nam cải tiến để có được kết quả tốt hơn.

Dự báo chuỗi thời gian là vấn đề luôn được nhiều nhà khoa học trên thế giới quan tâm nghiên cứu. Q.Song và B.S. Chissom [2] lần đầu tiên đã đưa ra quan niệm mới xem các giá trị thực định lượng trong chuỗi thời gian từ góc độ định tính. Từ đó chuỗi thời gian có thể xem như một biến ngôn ngữ và bài toán dự báo trở thành vấn đề dự báo các giá trị ngôn ngữ của biến ngôn ngữ. Có thể coi đây là quan niệm mới về chuỗi thời gian có tính đột phá. Tuy nhiên mô hình tính toán nhóm quan hệ mờ [5] quá phức tạp và do đó độ chính xác của dự báo không cao. Chen đã thay đổi cách tính toán nhóm quan hệ mờ trong mô hình dự báo [6, 7] với các phép tính số học đơn giản hơn để thu được kết quả dự báo chính xác hơn. Nhiều nghiên cứu tiếp theo vẫn sử dụng phương pháp luận này và đã thu được nhiều kết quả quan trọng.

Các nghiên cứu trên thế giới chủ yếu tập trung giải quyết vấn đề nâng cao độ chính xác dự báo. Có thể thấy một số vấn đề sau đây ảnh hưởng đến độ chính xác dự báo chuỗi thời gian mờ:

Mờ hóa các dữ liệu: Đây là vấn đề đòi hỏi phải có trực giác tốt để mô tả định tính chuỗi thời gian một cách hợp lý, từ đó xây dựng nhóm quan hệ mờ cung cấp thông tin có giá trị cho quá trình dự báo sau này. Đặc tính quan trọng của phép mờ hóa là số lượng khoảng chia, độ dài khoảng chia. Nếu số

lượng khoảng chia quá ít, dự báo có thể có độ sai lệch lớn do chưa đủ thông tin. Nếu số lượng khoảng chia quá lớn, dự báo có thể mất hết ý nghĩa về tính mờ của giá trị ngôn ngữ do không còn nhóm quan hệ mờ. Trong các nghiên cứu [10] số lượng khoảng, độ dài khoảng và bậc của mô hình chuỗi thời gian mờ có ảnh hưởng đến độ chính xác của mô hình dự báo. Một số nghiên cứu sâu hơn về số lượng khoảng, độ dài khoảng và bậc của mô hình chuỗi thời gian mờ tối ưu để có dự báo tốt nhất cho các dữ liệu trong nhóm quan hệ mờ.

Giải mờ: Đây là quá trình dự báo với rất nhiều kỹ thuật khác nhau trên cơ sở phép mờ hóa trên đây. Cách giải mờ phổ biến dựa trên 3 luật cơ bản [6], tuy nhiên trong một số tài liệu đã tìm ra một số tham số định hướng cho quá trình giải mờ và đã thu được một số kết quả khá tốt

Tiếp cận đại số gia tử (ĐSGT) [12] là tiếp cận khác biệt so với tiếp cận mờ và đã có một số ứng dụng thể hiện rõ hiệu quả của tiếp cận này so với tiếp cận mờ truyền thống trong một số lĩnh vực như điều, công nghệ thông tin. Tiếp tục những nghiên cứu ứng dụng trên đây, tiếp cận ĐSGT cũng cần được nghiên cứu thử nghiệm cho một lĩnh vực ứng dụng mới, đó là bài toán xây dựng mô hình dự báo chuỗi thời gian mờ đã được nhiều tác giả khác trên thế giới quan tâm hiện nay.

Đại số gia tử (ĐSGT) là một tiếp cận mới được các tác giả N.C.Ho và W. Wechler xây dựng vào những năm 1990, 1992 khi đưa ra một mô hình tính toán hoàn toàn khác biệt so với tiếp cận mờ. Những ứng dụng của tiếp cận ĐSGT cho một số bài toán cụ thể trong lĩnh vực công nghệ thông tin và điều khiển đã mang lại một số kết quả quan trọng khẳng định tính ưu việt của tiếp cận này so với tiếp cận mờ truyền thống.

Tuy nhiên, để lựa chọn bộ tham số tốt có thể phải cần đến nhiều lớp gia tử tác động lên phần tử sinh ban đầu trong biến ngôn ngữ và trên thực tế chỉ có nhiều nhất 3 lớp gia tử tác động. Vì vậy, nhiều giá trị ngôn ngữ trong biến ngôn